

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on:

December 8, 2001, 18:56:41 ; Search time 1735.73 Seconds
(without alignments)
24528.455 Million cell updates/sec

Title: US-08-153-397A-1

Perfect score: 3962

Sequence: 1 CGGCGCTGAGACTGGGCTGA.....AAAAAACCAGGAATTC 3962

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gdext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:*
- 1: em_estfun:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estom:*
 - 5: em_estpl:*
 - 6: em_estba:*
 - 7: em_estro:*
 - 8: em_estov:*
 - 9: em_hic:*
 - 10: gb_estl:*
 - 11: gb_est2:*
 - 12: gb_hic:*
 - 13: gb_gss:*
 - 14: em_gss_fun:*
 - 15: em_gss_hum:*
 - 16: em_gss_inv:*
 - 17: em_gss_pln:*
 - 18: em_gss_pro:*
 - 19: em_gss_rod:*
 - 20: em_gss_vrt:*
 - 21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1720	43.4	2633	12	BC006836	Mus muscu
2	960	24.2	1010	10	AL539517	AL539517 AL539517
3	815.2	20.6	1019	10	AL528663	AL528663 AL528663
4	814.6	20.6	987	10	AL528664	AL528664 AL528664
5	715.8	18.1	770	11	BI257820	BI257820 60296918
6	684	17.3	799	11	BG12321	BG12321 602806864
7	580.4	17.2	715	10	BE74806	BE74806 601575025
8	672.4	17.0	719	11	BF345815	BF345815 602017891
9	672	17.0	744	11	BI193181	BI193181 602947157
10	662.6	16.7	718	11	BG696171	BG696171 602659285
11	660.2	16.7	772	10	AL043251	AL043251 DKF2p434K
12	655.6	16.5	912	10	BE304984	BE304984 600942930

13	654.6	16.5	652	11	BF944646
14	653.8	16.5	877	10	BE249888
15	652	16.5	957	10	BE737252
16	650.8	16.4	683	11	BE752521
17	650.4	16.4	685	11	BG403848
18	646.4	16.3	787	11	BF338761
19	646	16.3	1144	11	BF527983
20	643.6	16.2	898	11	BE898332
21	640.8	16.2	801	10	BE735635
22	639.6	16.1	1018	11	BE899403
23	636.2	16.1	759	11	BF344637
24	635	16.0	921	11	BF792714
25	632.2	16.0	1009	10	BE613181
26	631.2	15.9	759	11	BG913935
27	629.2	15.9	711	11	BG541910
28	628.4	15.9	1150	11	BF344537
29	627.4	15.8	652	10	AL047539
30	627.2	15.8	804	11	BG696424
31	626.6	15.8	977	11	BG386352
32	624	15.7	732	11	BF339561
33	621	15.7	831	10	BE513178
34	620.8	15.7	626	10	AW376875
35	619.6	15.6	1037	11	BF29368
36	613	15.5	680	11	BG818241
37	603.6	15.2	649	10	BE304585
38	602.4	15.2	1325	11	BF966734
39	594.8	15.0	645	10	BG395129
40	591.2	14.9	904	11	BG289311
41	588.4	14.9	614	10	BE305175
42	586.8	14.8	600	10	AW629699
43	583.8	14.7	611	11	BG820294
44	583	14.7	596	11	BG831101
45	580.8	14.7	922	11	BG116520

ALIGNMENTS

RESULT 1

BC006836

LOCUS

BC006836 2633 bp mRNA

Mus musculus, Similar to discoidin domain receptor family, member

1, clone IMAGE:3598142, mRNA.

ACCESSION BC006836

VERSION BC006836.1

KEYWORDS GI:14711754

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2633)

Strausberg, R.

Direct Submission

Submitted (27-APR-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: villalona@bcm.tmc.edu

Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,

A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,

Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

HTC

12-JUL-2001

member


```
QY 3040 TGCATCGGTCTCTGGCAGAGGATGCACATCAACACGGTGTGAATCACACATCCAGCTGCC 3099
Db 539 TGCATCGGTCTCTGGCAGAGGATGCACATCAACACGGTGTGAATCACACATCCAGCTGCC 598
QY 3100 CTCCTCAGGAGGTATCCAGGGAAGCCAGTGCACACTAAACAAAGAGGACACAAATGGCA 3159
Db 599 CTCCTCAGGAGGTATCCAGGGAAGCCAGTGCACACTAAACAAAGAGGACACAAATGGCA 658
QY 3160 CCTCTGCCCTTCCCTCCCGACAGCCATCCTCTTAATAGAGGAGTGCAGGT 3219
Db 659 CCTCTGCCCTTCCCTCCCGACAGCCATCCTCTTAATAGAGGAGTGCAGGT 718
QY 3220 GGGTGGGCCCAACCCAGGAGCTGATGCCCTTCTCCCTTCTCCGACACACTCTCATGT 3279
Db 719 GGGTGGGCCCAACCCAGGAGCTGATGCCCTTCTCCCTTCTCCGACACACTCTCATGT 778
QY 3280 CCCCTCTCTGTTCTTCTTGAAGCCCTGTCGCCCAACCCAGTGGTCTCTGGATG 3339
Db 779 CCCCTCTCTGTTCTTCTTGAAGCCCTGTCGCCCAACCCAGTGGTCTCTGGATG 838
QY 3340 GGATCCTCTCCACCTCTCTAGCCATCCCTTGGGGAAGGTTGGGGAAGAAATAGGATA 3399
Db 839 GGATCCTCTCCACCTCTCTAGCCATCCCTTGGGGAAGGTTGGGGAAGAAATAGGATA 898
QY 3400 GACACTGGACATGCCCATTTGGACACCTGGGCCCACTGGACACACTGATTCCTGGAG 3459
Db 899 GACACTGGACATGCCCATTTGGACACCTGGGCCCACTGGACACACTGATTCCTGGAG 958
QY 3460 AGGTGGTGGGCCCCAGCTTCTCTCTCC 3487
Db 959 A-GKGGTGGGCCCCAGCTTCTCTCTCC 985

RESULT 3
AL528663/c
LOCUS
DEFINITION AL528663 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD001YK21 3
prime, mRNA sequence.
ACCESSION AL528663
VERSION AL528663.1 GI:12792156
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 1019
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DD001YK21"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
299 a 272 c 242 g 179 t 27 others
```

ORIGIN

Query Match 20.6%; Score 815.2; DB 10; Length 1019;
Best Local Similarity 89.08; Pred. No. 3.9e-141;
Matches 906; Conservative 22; Mismatches 85; Indels 5; Gaps 5;

```
QY 2893 AGGTCAATCGAAGACCGGGGGAGTCTTCCGGGACAGCGCGGAGGTGTACCTGTCC 2952
Db 1016 AAGTCATCGAAGACCGGGGGAKTT-TTCCGGGACAGCGCGGAGGTGTACCTGTCC 958
QY 2953 GGGCGCCTGCTCCCGCAGGCGCTATATGAGTGTCTCGGTGCTGGAGCCGGAGT 3012
Db 957 GGCNGCCTGCGCTG-CCGAAGGGCATATATGAGTGTCTCGGTGCTGGAGCCGGAGT 899
QY 3013 CTGAGCAGCGACACACCTTTTCCAGCTGCATCGGTTCCTGCACAGAGATGCACACA 3072
Db 898 YTGAGCAGCGACACACCTTTTCCRGCTGCATCGGTTCCTGCAGAGAGATGCACACA 839
QY 3073 CGGTGTGAATCACACATCCAGCTCCGCTCCTCAGGAGTGCATCCAGGGGAAGCCAGTG 3132
Db 838 CGGTGTGAATCACACATCCAGCTCCGCTCCTCAGGAGCGATCCAGGGGAAGCCAGTG 779
QY 3133 ACATAAACAAGAGGACACAAATGGCCTCTCCCTTCCCTCCCGACAGCCCATACC 3192
Db 778 ACATAAACAAGAGGACACAAATGGCCTCTCCCTTCCCTCCCGACAGCCCATACC 719
QY 3193 TCTAATAGAGCAGTGCAGCTG-CAGGTGGCTGGGCCACCCAGGGAGCTGTGCCCT 3251
Db 718 TCTAATAGAGCAGTGCAGCTGCCAGGTGGCTGGGCCACCCAGGGAGCTGTGCCCT 659
QY 3252 TCTCCCTTCTCTGGACACACTCTCATGTCTCCCTTCTCTTCTCTTCTAGAACCTCT 3311
Db 658 TCTCCCTTCTCTGRACAMACTCTCATGTCCCTTCTCTTCTCTTCTAGAACCTCT 599
QY 3312 GTGCCCAACAGCTGTCT-CTGTGGATGGATCCTCTCCACCCTCTCTTACCATCCT 3370
Db 598 GTGCCCTACCTGTGTGTCTGTGGATGGATTTCTCCACCTTCTTTTACCATCCTT 539
QY 3371 TGGGGAAGGTGGGGAAGAAATATAGATAGACACTGCACATGGCCCATTTGAGACCTGG 3430
Db 538 TGGGGAAGGTGGGGAAGAAATATAGATAGATAGATAGATAGATAGATAGATAGATAG 479
QY 3431 GCGCCACTGGACACACACTGATTCCTGGAGAGTGGCTGCGCCCGCAGCTTCTCTCCCTG 3490
Db 478 GCGCTATTGGACTATATTGATTCTGGAGAGGTGTGAGC-CCCTGTTTTTTTTCTGTG 420
QY 3491 TCACACACTGGACCCCACTCGCTGAGAAATCTGGGGGTGAGGAGACAAAGAGGAGGAA 3550
Db 419 TTACATATTGGACCTATTGGCTGAGAAATTTGGGGGTGAGGAGACAAAGAGGAGGAA 360
QY 3551 AATGTTTCTTGTGCTGCTCTCTGTTCTGCTCAGCTGGGGTCTTCTCTCTCTCATC 3610
Db 359 AATGTTTCTTGTGCTGCTCTCTGTTCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 3611 ACCTGAACACTGGACCTGGGGTAGCCCGCCCGCCAGCTCAGTCAACCCCACTTCCCA 3670
Db 299 ACCTGAACACTGGACCTGGGGTAGCCCGCCCGCCAGCTTCTGTTACCTTCTCTTA 240
QY 3671 CTTCAGTCTTGTAGTAGAACTTCTTAAGCCTATACGTTTCTGTGGAGTAATATTGG 3730
Db 239 CTTCAGTCTTGTAGTAGAACTTCTTAAGCCTATACGTTTCTGTGGAGTAATATTGG 180
QY 3731 GATTGGGGGAAGAGGAGGACACAGCCCACTTGGGGTGGGAGATCTCTAGTGTAG 3790
Db 179 GATTGGGGGAAGAGGAGGACACAGCCCACTTGGGGTGGGAGTCTTCTAGTGTAG 120
QY 3791 CTGCCACATGATTTTCTTATATCACTTGGGGTGTGTACATTTTGGGGGGGAGACAC 3850
Db 119 CTGCCACATGATTTTCTTATATCACTTGGGGTGTGTACATTTTGGGGGGGAGACAC 60
QY 3851 AGATTTTACACTAATATATGACCTAGCTTGGAGCAATTTTAAATCCCTGCACTAGG 3908
Db 119 AGATTTTACACTAATATATGACCTAGCTTGGAGCAATTTTAAATCCCTGCACTAGG 3908
```

Db 59 TGATTTTACACTTATATTGGTCCCTGCTGTGAGGAAATTTTAATATATGAWTTAGG 2

RESULT 4
AL528664 987 bp mRNA EST 13-FEB-2001
LOCUS LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD001YK21 5
DEFINITION prime, mRNA sequence.
AL528664
ACCESSION AL528664
VERSION
KEYWORDS
SOURCE EST
GI:12792157
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 987)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1. .987
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DD001YK21"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 186 a 263 c 332 g 200 t 6 others
BASE ORIGIN

Query Match 20.6%; Score 814.6; DB 10; Length 987;
Best Local Similarity 98.7%; Pred. No. 5e-141;
Matches 848; Conservative 4; Mismatches 3; Indels. 4; Gaps 3;

QY 277 AGAGATGCTGCCCCACCCCTTAGCCCGAGGATCAGAGGTATGGACACAGAGGCC 336
Db 90 AGARATGCTGCCCCACCCCTTAGCCCGAGGATCAGAGGTATGGACACAGAGGCC 149
QY 337 TGTATCTTTACTGCTGCTCTTGGTGGCAAGTGCATGATGAAGGACATT 396
Db 150 TGTATCTTTACTGCTGCTCTTGGTGGCAAGTGCATGATGAAGGACATT 209
QY 397 TTGATCTCCCAAGTGCCTATGCTGGCATGACGACCGGACCATCCACAGAGTG 456
Db 210 TTGATCTCCCAAGTGCCTATGCTGGCATGACGACCGGACCATCCACAGAGTG 269
QY 457 ACATCTCTCTTCCAGTCTCTGTCAGATCCACTGCCCGCCGCCACACAGGTTGGAGA 516
Db 270 ACATCTCTCTTCCAGTCTCTGTCAGATCCACTGCCCGCCGCCACACAGGTTGGAGA 329
QY 517 GCAGTGACGGGATGGGGCTGGTCCCGCAGGCTGGTGTTCCTCCAGGAGGAGT 576
Db 330 GCAGTGACGGGATGGGGCTGGTCCCGCAGGCTGGTGTTCCTCCAGGAGGAGT 387
QY 577 ACTTCGAGTGATGTACACGACTCCACTCTGGTGGCTCTGGTGGGACCCACGAGGACGC 636
Db 388 ACTTCGAGTGATGTACACGACTCCACTCTGGTGGCTCTGGTGGGACCCACGAGGACGC 447

QY 637 ATGCGGGGGCTGGGCAAGAGTTCTCCCGAGCTACCGGCTGGTGTACTCCCGGGATG 696
Db 448 ATGCGGGGGCTGGGCAAGAGTTCTCCCGAGCTACCGGCTGGTGTACTCCCGGGATG 507
QY 697 GTCCCGCTGATGGCTGGGAGGACCGCTGGGTCAGAGGTGATCTCAGGCAATGAGG 756
Db 508 GTCCCGCTGATGGCTGGGAGGACCGCTGGGTCAGAGGTGATCTCAGGCAATGAGG 567
QY 757 ACCGTAGGGAGTGTGCTGAAGGACCTTGGGCCCCCATGGTGGCCGACTGTTCCGT 816
Db 568 ACCGTAGGGAGTGTGCTGAAGGACCTTGGGCCCCCATGGTGGCCGACTGTTCCGT 627
QY 817 TCTACCCCGGGCTGACCGGGTCTAGTGTCTGTCTGGGGTAGAGCTCTATGGCTGCC 876
Db 628 TCTACCCCGGGCTGACCGGGTCTAGTGTCTGTCTGGGGTAGAGCTCTATGGCTGCC 687
QY 877 TCTGAGGGATGGACTCTCTGTCTTACACCGCCCTGTGGGCGACACATGTTATCTG 936
Db 688 TCTGAGGGATGGACTCTCTGTCTTACACCGCCCTGTGGGCGACACATGTTATCTG 746
QY 937 AGCCGTGTACTCAACGACTCCACCTATGACGACATACCGTGGGGGACTGCAGTATG 996
Db 747 AGCCGTGTACTCAACGACTCCACCTATGACGACATACCGTGGGGGACTGCAGTATG 806
QY 997 GGGTCTGGCCAGCTGGCAGATGGTGTGGGGCTGGATGACTTTAGGAAGAGTCAGG 1056
Db 807 GGGTCTGGCCARCTGGCAGATGGTGTGGGGCTGGATGACTTTAGGAAGAGTCAGG 866
QY 1057 AGCTGGGGTCTGGCCAGCTATGACTATGAGTGGGATGGAGCAACACAGCTTCCAGTG 1116
Db 867 AGCTGGGGTCTGGCCAGCTATGACTATGAGTGGGATGGAGCAACACAGCTTCCAGTG 925
QY 1117 GCTATGTGAGATGGAGTT 1135
Db 926 GCTATGTGAGATGGAGTT 944

RESULT 5
BI257820 770 bp mRNA EST 17-JUL-2001
LOCUS 602969918F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5109436 5',
DEFINITION mRNA sequence.
ACCESSION BI257820.1 GI:14813571
VERSION BI257820
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 770)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11265 row: c column: 05
High quality sequence stop: 765.
Location/Qualifiers
1. .770
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5109436"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"

FEATURES
source

/note="Organ: cervix; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
144 a 207 c 262 g 157 t

BASE COUNT 144 a 207 c 262 g 157 t

ORIGIN

Query Match 18.1%; Score 715.8; DB 11; Length 770;
Best Local Similarity 97.3%; Pred. No. 1e-122;
Matches 750; Conservative 0; Mismatches 17; Indels 4; Gaps 2;

QY 229 GAAGTGGCTATTCACTAGCGATGGGTGGACTTGAAGGATGCCAAGAGATGCTGCC 288
|||||
Db 1 GAAGTGGCTATTCACTAGCGATGGGTGGACTTGAAGGATGCCAAGAGATGCTGCC 60
QY 289 CCACACCCCTTAGGCCCGGAGGATGAGGATGATGGACAGAGGCCCTGTCTATCTTAC 348
|||||
Db 61 CCACACCCCTTAGGCCCGGAGGATGAGGATGATGGACAGAGGCCCTGTCTATCTTAC 120
QY 349 TGGTGTCTCTTGTGGCAAGTGGAGATGCTGACATGAAGGACATTTTGTATCTTGCCA 408
|||||
Db 121 TGGTGTCTCTTGTGGCAAGTGGAGATGCTGACATGAAGGACATTTTGTATCTTGCCA 180
QY 409 AGTGGCGCTATGCCCTGGGCATGCGAGGACCGGACCATCCACAGCATGACATCTTGCTT 468
|||||
Db 181 AGTGGCGCTATGCCCTGGGCATGCGAGGACCGGACCATCCACAGCATGACATCTTGCTT 240
QY 469 CCAGCTCTCTGGTCAGATTCCTACCTGCCGCCGCCACAGCAGGTGGAGACATGACGGGG 528
|||||
Db 241 CCAGCTCTCTGGTCAGATTCCTACCTGCCGCCGCCACAGCAGGTGGAGACATGACGGGG 300
QY 529 ATGGGCGCTTGGTGGCCCGCAGGCTGGGTGTTTCCCAAGGAGGAGGATGACTTGCAGGTGG 588
|||||
Db 301 ATGGGCGCTTGGTGGCCCGCAGGCTGGGTGTTTCCCAAGGAGGAGGATGACTTGCAGGTGG 360
QY 589 ATCTACAGCATCCACCTGCTGGTCTGGTGGGACCCAGGAGCGCATGCCGGGGGCC 648
|||||
Db 361 ATCTACAGCATCCACCTGCTGGTCTGGTGGGACCCAGGAGCGCATGCCGGGGGCC 420
QY 649 TGGCAAGGAGTCTCCCGGAGCTACCGGCTGCGTTACTCCCGGATGGTCCCGCTGGA 708
|||||
Db 421 TGGGCAAGGAGTCTCCCGGAGCTACCGGCTGCGTTACTCCCGGATGGTCCCGCTGGA 480
QY 709 TGGGCTGGAAGGACCGCTGGGTTCAGGAGGTGATCTCAGGCAATGAGGACCTGAGGGAG 768
|||||
Db 481 TGGGCTGGAAGGACCGCTGGGTTCAGGAGGTGATCTCAGGCAATGAGGACCTGAGGGAG 540
QY 769 TGGTGTCTGAAGGACCTTGGGCCCCCCTGTTGGCCGAGTGTGCTTCTACCCCGGG 828
|||||
Db 541 TGGTGTCTGAAGGACCTTGGGAAACCCCTGTTGGCCGAGTGTGCTTCTACCCCGGG 600
QY 829 CTGACCGGTCTATG---AGTGTCTCTGCGGGTAGAGCTCTATGGCTGCTCTGAGGG 885
|||||
Db 601 CTGACCGGTCTATGAGCAGTCTGTACTGCGGGTAGAAGCTCTATGGCTGCTCTGAGGG 660
QY 886 ATGGACTCTCTGTCTACACCCCTGTTGGGCGCAGACATGATTTATCTGAGGCCGTGT 945
|||||
Db 661 ATGGACTCTCTGTCTACACCCCTGTTGGGCGCAGACATGTA-TTATCTGAGGCCGTGT 719
QY 946 ACCTACAGCATCCACCTATGACGACATACCGTGGGGGGGACTGACGATG 996
|||||
Db 720 ACCTACAGCATCCACCTATGACGACATACCGTGGGGGGGACTGACGATG 770

RESULT 6
BG912321
LOCUS BG912321 799 bp mRNA EST 05-JUN-2001
DEFINITION 602806864F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4939239
5', mRNA sequence.
ACCESSION BG912321
VERSION BG912321.1 GI:14292797
KEYWORDS EST.

SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 799)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10875 row: o column: 16
High quality sequence stop: 768.
Location/Qualifiers
1...799
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4939239"
/clone_lib="NCI_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with lp/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 163 a 238 c 251 g 147 t
ORIGIN

Query Match 17.3%; Score 684; DB 11; Length 799;
Best Local Similarity 95.7%; Pred. No. 7.8e-117;
Matches 758; Conservative 0; Mismatches 25; Indels 9; Gaps 5;

QY 2443 ACTACATGAGAACGGCGACCTCAACCACTTCTCTAGTGGCCACGAGCTGGAGGACAAG 2502
|||||
Db 1 ACTACATGAGAACGGCGACCTCAACCACTTCTCTAGTGGCCACGAGCTGGAGGACAAG 60
QY 2503 CAGCCGAGGGGGCCCTTGGGGACGGCGAGCTGCGCAGGGGCCACCATCAGCTACCCAA 2562
|||||
Db 61 CAGCCGAGGGGGCCCTTGGGGACGGCGAGCTGCGCAGGGGCCACCATCAGCTACCCAA 120
QY 2563 TGCTGCTGATGTGGCAGCCCGAGATGCGCTCCGGCATGCGCTATCTGGCCACACTCAACT 2622
Db 121 TGCTGCTGATGTGGCAGCCCGAGATGCGCTCCGGCATGCGCTATCTGGCCACACTCAACT 180
QY 2623 TTGTACATCGGACCTGGCCAGCGGAACTGCTAGTGGGGAATTTTCACCATCAAA 2682
|||||
Db 181 TTGTACATCGGACCTGGCCAGCGGAACTGCTAGTGGGGAATTTTCACCATCAAA 240
QY 2683 TGGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGAGCTATTACCGTGTGAGGGCC 2742
|||||
Db 241 TGGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGAGCTATTACCGTGTGAGGGCC 300
QY 2743 GGGCAGTGTGCTCCCATCCGCTGGATGGCTGGGAGTGCATCTCATGGGGAAGTTCACGA 2802
Db 301 GGGCAGTGTGCTCCCATCCGCTGGATGGCTGGGAGTGCATCTCATGGGGAAGTTCACGA 360
QY 2803 CTGGCAGTGTGCTGGGCTTTGGTGACCCCTTGGTGGGAGGTGCTGATGCTGTAGGG 2862
|||||
Db 361 CTGGCAGTGTGCTGGGCTTTGGTGACCCCTTGGTGGGAGGTGCTGATGCTGTAGGG 420
QY 2863 CCCAGCCCTTTGGGAGCTCACCGAGCAGCAGTGTATCCAGACGCGGGGAGTTCCTTC 2922
Db 421 CCCAGCCCTTTGGGAGCTCACCGAGCAGCAGTGTATCCAGACGCGGGGAGTTCCTTC 480
QY 2923 GGGACAGGGCGGAGGTGTACTGTCCCGCCGCCCTGCCTGCCCGC-AGGSCCTATAT 2981

Query Match 17.2%; Score 680.4; DB 10; Length 715;
Best Local Similarity 99.4%; Pred. No. 3.6e-116;
Matches 714; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

RESULT	8
BF345815	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LAM9421 row: f column: 24
High quality sequence stop: 719.
Location/Qualifiers

FEATURES

Source

1..719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4153367"
/clone_lib="NCI_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with lp/19q loss"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 132 a 217 c 230 g 140 t
ORIGIN

Query Match 17.0%; Score 672.4; DB 11; Length 719;
Best Local Similarity 99.3%; Pred. No. 1.1e-114;
Matches 717; Conservative 0; Mismatches 1; Indels 4; Gaps 4;
QY 2397 CTGCTGGCGGTGTGTGTCAGGAGACCCCTCTGCATGATTACTACTACATGAGAAC 2456
Db 1 CTGCTGGCGGTGTGTGTCAGGAGACCCCTCTGCATGATTACTACTACATGAGAAC 60
QY 2457 GCGACCTCAACAGTTCTCAGTGCACCCAGCAGTGGAGGACAAAGGAGCGGAGGGGCC 2516
Db 61 GCGACCTCAACAGTTCTCAGTGCACCCAGCAGTGGAGGACAAAGGAGCGGAGGGGCC 120
QY 2517 CTTGGGAGCGGAGGTGCGAGGCGCCACCATCATGACCTACCAATGCTGCTCATGTG 2576
Db 121 CTTGGGAGCGGAGGTGCGAGGCGCCACCATCATGACCTACCAATGCTGCTCATGTG 180
QY 2577 GCAGCCAGATCGCTCGGATCGCTATCTGGCCACATCACTTTGTACATCGGGAC 2636
Db 181 GCAGCCAGATCGCTCGGATCGCTATCTGGCCACATCACTTTGTACATCGGGAC 240
QY 2637 CTGCGCCAGCGGAACTGCTAGTTGGGAAATTTACACATCAAAATTCAGACTTTGGC 2696
Db 241 CTGCGCCAGCGGAACTGCTAGTTGGGAAATTTACACATCAAAATTCAGACTTTGGC 300
QY 2697 ATGAGCGGAACTTATGCTGGGACTATTACGTTGTCAGGCGCGGAGTCTGTCGCC 2756
Db 301 ATGAGCGGAACTTATGCTGGGACTATTACGTTGTCAGGCGCGGAGTCTGTCGCC 360
QY 2757 ATCCGCTGGATGCGCTGGGAGTGCATCTCATGGGAAGTTACGACTGCGAGTGCAGTG 2816
Db 361 ATCCGCTGGATGCGCTGGGAGTGCATCTCATGGGAAGTTACGACTGCGAGTGCAGTG 420
QY 2817 TGGGCTTTGGTGTGACCTGTGGAGTGTCTGATGCTGTAGGCGCCAGCCCTTTGGG 2876
Db 421 TGGGCTTTGGTGTGACCTGTGGAGTGTCTGATGCTGTAGGCGCCAGCCCTTTGGG 480
QY 2877 CAGCTACCCAGCAGGATCATCGAGACGCGGGAGTTCTTCGGGACGAGGCGCG 2936
Db 481 CAGCTACCCAGCAGGATCATCGAGACGCGGGAGTTCTTCGGGACGAGGCGCG 539
QY 2937 CAGGTGTACCTGTCCGGCGCCCTGCTGCCCGCAGCGGCTATATGAGTGTGCTTCGG 2996
Db 540 CAGGTGTACCTGTCCGGCGCCCTGCTGCCCGCAGCGGCTATATGAGTGTGCTTCGG 599
QY 2997 TGTGGAGCGGAGTGTGAGCAGCAGCACCCCTTTTCCAGTGTGATCGGTTCTTGGA 3056
Db 600 TGTGGAGCGGAGTGTGAGCAGCAGCACCCCTTTTCCAGTGTGATCGGTTCTTGGA 658
QY 3057 GAGGATGCACTCAACA-CGGTGTGATCAGATCATCATGCTGCCCTCCCTCAGGGAGTGA 3115
Db 659 GAGGATGCACTCAACA-CGGTGTGATCAGATCATCATGCTGCCCTCCCTCAGGGAGTGA 717

OY 3116 TC 3117
Db 718 TC 719

RESULT 9

BI193181
LOCUS BI193181 744 bp mRNA EST 10-JUL-2001
DEFINITION 602947157F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5090223 5', mRNA sequence.
ACCESSION BI193181
VERSION BI193181.1 GI:14648201
KEYWORDS EST
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 744)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Cloning by: Incyte Genomics, Inc.
Cloned distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM852 row: b column: 16
High quality sequence stop: 675.
Location/Qualifiers
1..744
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5090223"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. |"
BASE COUNT 149 a 212 c 229 g 154 t
ORIGIN

Query Match 17.0%; Score 672; DB 11; Length 744;
Best Local Similarity 97.5%; Pred. No. 1.3e-114;
Matches 714; Conservative 0; Mismatches 15; Indels 3; Gaps 3;

QY 2237 TTTCCCTTAATGTGCGTAAGGACACCCCTTTGCTGCTAGCTGTCAAGATCTTACGGCC 2296
Db 2 TTTCCCTTAATGTGCGTAAGGACACCCCTTTGCTGCTAGCTGTCAAGATCTTACGGCC 61
QY 2297 AGATGCCACCAAGATGCCAGCTTCTCTTCTCCAGGAATGATTCTCTGAAGAGGT 2356
Db 62 AGATGCCACCAAGATGCCAGCTTCTCTTCTCCAGGAATGATTCTCTGAAGAGGT 121
QY 2357 GAAGATCATGTCCAGCTCAAGGACCCCAACATCATTCGGCTGCTGGCGTGTGTGCA 2416
Db 122 GAAGATCATGTCCAGCTCAAGGACCCCAACATCATTCGGCTGCTGGCGTGTGTGCA 181
QY 2417 GGACGACCCCTCTGCTGATGATTACTGACTACATGAGAGCGGAGCTCAACAGTCTCT 2476
Db 182 GGACGACCCCTCTGCTGATGATTACTGACTACATGAGAGCGGAGCTCAACAGTCTCT 241

TITLE EST (Blum, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Blum H
MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DFPZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.

No s1 sequence available.

This clone (DKF2p43K0223) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

Location/Qualifiers

1..772

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="DKF2p43K0223"

/clone_lib="434 (synonym: htes3)"

/tissue_type="testis"

/dev_stage="adult"

/lab_host="DH10B"

/note="vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 118 a 250 c 260 g 143 t

ORIGIN

Query Match

Best Local Similarity 16.7%; Score 660.2; DB 10; Length 772;

Matches 750; Conservative 0; Mismatches 4; Indels 58; Gaps 3;

QY 64 CTCCTGGAGCGCTCCCGACACCCGAGCGCGCTCCCGCTCCCGCTCCCGCTCCCG 123

DB 1 CTCCTGGAGCGCTCCCGACACCCGAGCGCGCTCCCGCTCCCGCTCCCGCTCCCG 60

QY 124 GCTCTGGCTCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 183

DB 61 GCTCTGGCTCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 120

QY 184 GGTCTGGAGCGCTGGTCTCCGCGGAGAGCGATGAGAGGTGCTGAAGTGGCTATCA 243

DB 121 GGTCTGGAGCGCTGGTCTCCGCGGAGAGCGATGAGAGGTGCTGAAGTGGCTATCA 157

QY 244 CTGAGCGATGGGTGGACTTGAAGGAATGCCAAGAGATGTCGCCCGACCCCTTAGGC 303

DB 158 -----AGATGCTGCCCGACCCCTTAGGC 184

QY 304 CCGAGGATCAGGAGCTATGGACCGAGCGCTGCTCTTACTGCTGCTGCTGCTG 363

DB 185 CCGAGGATCAGGAGCTATGGACCGAGCGCTGCTCTTACTGCTGCTGCTGCTG 244

QY 364 TGGCAAGTGGAGTGTGACATGAAGGACATTTGATCTGCAAGTGGCGCTATGCC 423

DB 245 TGGCAAGTGGAGTGTGACATGAAGGACATTTGATCTGCAAGTGGCGCTATGCC 304

QY 424 TGGCATGCAGGACCGGACCATCCAGACAGTGCATCTGCTTCCAGCTCCTGGTGCAG 483

DB 305 TGGCATGCAGGACCGGACCATCCAGACAGTGCATCTGCTTCCAGCTCCTGGTGCAG 364

QY 484 ATTCCACTGGCGCCCGACAGAGGTTGAGAGCAGTGCAGGGGATGGGGCTGGTGC 543

DB 365 ATTCCACTGGCGCCCGACAGAGGTTGAGAGCAGTGCAGGGGATGGGGCTGGTGC 424

QY 544 CCGCAGGCTGGTGTTCACAGGAGGAGTACTTGCAGGTGGATCTACACAGCTCC 603

DB 425 CCGCAGGCTGGTGTTCACAGGAGGAGTACTTGCAGGTGGATCTACACAGCTCC 484

QY 604 ACCTGTGGCTGTGGTGGGACCCAGGCGGATGCCGGGGCTGGGCAAGGAGTTCT 663

DB 485 ACCTGTGGCTGTGGTGGGACCCAGGCGGATGCCGGGGCTGGGCAAGGAGTTCT 544

QY 664 CCGGAGCTACCGGTGCGTTACTCCCGGATGTCGCGCTGGTGGCTGAAGGACC 723

DB 545 CCGGAGCTACCGGTGCGTTACTCCCGGATGTCGCGCTGGTGGCTGAAGGACC 604

QY 724 GCTGGGTTCAGGAGTGTCTCAGGCATGAGGACCCCTGAGGAGTGGTCTGAAGGACC 783

DB 605 GCTGGGTTCAGGAGTGTCTCAGGCATGAGGACCCCTGAGGAGTGGTCTGAAGGACC 664

QY 784 TTGGCCCCCATGTTCCCGAGCTGGTTCCTTACCCCGGCTGACCCGGTTCATGA 843

DB 665 TTGG-CCCCCATGTTCCCGAGCTGGTTCCTTACCCCGGCTGACCCGGTTCATGA 723

QY 844 GTGCTGTCTCCGGGTAGAGCTCTATGGCTGC 875

DB 724 GCGTCTGTCTCCGGGTAGA-CTCTATGGCTGC 754

RESULT 12

BE304984/c

LOCUS

DEFINITION

600942930T2 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2959615 3',

mRNA sequence.

ACCESSION

BE304984

VERSION

BE304984.1 GI:9176727

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 912)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: rgs@bbs-rcmail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: NIH Intramural Sequencing Center

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LCM51 row: k column: 08

High quality sequence start: 9

High quality sequence stop: 702.

FEATURES

Location/Qualifiers

1..912

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2959615"

/clone_lib="NIH_MGC_15"

/tissue_type="adenocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 211 a 244 c 277 g 180 t

ORIGIN

Query Match

Best Local Similarity 16.5%; Score 655.6; DB 10; Length 912;

Matches 740; Conservative 0; Mismatches 29; Indels 7; Gaps 6;

QY 3098 CCTCTCCTCAGGAGTGTATCCAGGGAGCCAGTGCAC--TAAACAAGAGGACAAAT 3155

DB 778 CCTCTCCTCAGGAGTGTATCCAGGGAGCCAGTGCAC--TAAACAAGAGGACAAAT 719.

QY 3156 GGCACCTCTGCCCTTCCCTCCCGACAGCCATCACCTCTTAATAGAGGAGTGC 3215


```

|||||
Db 601 TAGACACTGGACATGGGCGCATTTGGACACCTTGGGCCCGCCACTGGACACACTGATTCCTGG 660
Qy 3458 AG 3459
Db 661 AG 662

RESULT 14
LOCUS BE249888 877 bp mRNA EST 13-JUL-2000
DEFINITION 600942930F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2959615 5',
mRNA sequence.
ACCESSION BE249888
VERSION BE249888.1 GI:91119991
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 877)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: LLCM51 row: k column: 08
High quality sequence stop: 725.
FEATURES
Location/Qualifiers
1..877
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2959615"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 167 a 248 c 282 g 180 t
ORIGIN

Query Match 16.5%; Score 653.8; DB 10; Length 877;
Best Local Similarity 91.0%; Pred. No. 3e-111;
Matches 782; Conservative 0; Mismatches 47; Indels 30; Gaps 7;

Qy 2082 GTGCGTGCACCTGCCCGGCGAGTGGGATGGGCGGCCCGCCAGAGTGGATTCCCTCGA 2141
Db 1 GTGCGTGCACCTGCCCGGCGAGTGGGATGGGCGGCCCGCCAGAGTGGATTCCCTCGA 60

Qy 2142 TCTCGACTCCGCTCAAGGAGAGCTTGGCGAGGCGCCAGTTGGGGAGGTGCACCTGTGT 2201
Db 61 TCTCGACTCCGCTCAAGGAGAGCTTGGCGAGGCGCCAGTTGGGGAGGTGCACCTGTGT 120

Qy 2202 GAGGTGCGACAGCCCTCAAGATCTGTGATGTC-TTGATTCCCGCTTAATGTGCGTAGGG 2260
Db 121 GAGGTGCGACAGCCCTCAAGATCTGTGATGTC-TTGATTCCCGCTTAATGTGCGTAGGG 180

Qy 2261 ACACCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2320
Db 181 ACACCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 233

```

```

Qy 2321 CTCCTTGTTCTCCAGGAATGATTTCTTGAAGAGGTGAAGATCATCTGAGGCTCAAGGA 2380
Db 234 -----CCAGGAATGATTTCTTGAAGAGGTGAAGATCATCTGAGGCTCAAGGA 282
Qy 2381 CCCCAACATCATTCGGCTGCTGGGCGTGTGTGTGAGGACGACCCCTCTGATGATTAC 2440
Db 283 CCCAAACATCATTCGGCTGCTGGGCGTGTGTGTGAGGACGACCCCTCTGATGATTAC 342
Qy 2441 TGACTACATGAGAACGGCGACCTCAACCACTTCTCTAGTCCCGACAGCTGGAGGACAA 2500
Db 343 TGACTACATGAGAACGGCGACCTCAACCACTTCTCTAGTCCCGACAGCTGGAGGACAA 402
Qy 2501 GGCAGCGAGGGGGCCCTTGGGAGCGGCGAGGCTGCGCAGGGGCGCCACCATCAGCTACCC 2560
Db 403 GGCAGCGAGGGGGCCCTTGGGAGCGGCGAGGCTGCGCAGGGGCGCCACCATCAGCTACCC 462
Qy 2561 AATGCTCTGCATGTGGCAGCCCGCATGCGCTCCGGCATGCGCTATCTGGCCACACTCAA 2620
Db 463 AATGCTCTGCATGTGGCAGCCCGCATGCGCTCCGGCATGCGCTATCTGGCCACACTCAA 522
Qy 2621 CTTTGTACATCGGGACCTGGCCACGCGAACTGCTAGTTGGGAAATTTTCCACATCAA 2680
Db 523 CTTTGTACATCGGGACCTGGCCACGCGAACTGCTAGTTGGGAAATTTTCCACATCAA 582
Qy 2681 AATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTACCGTGTGCA--G 2738
Db 583 AATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTACCGTGTGCACTG 642
Qy 2739 GCGCGGCGATGTCGCCATCGCTGCGCTGGAGTGGAGTGCATCTCATGGG-AACTT 2797
Db 643 GTCGCGAGTGTGTCCTATCGCTGGGATGGCTGGGAATGCTCCCTCATGGGAAAGTT 702
Qy 2798 CAGGACTGCGAGTACGTGTGGCCCTTTGGTGTGACCTCTGGGAGGTGCTGATGCTGTG 2857
Db 703 CAGGACTGCGAGTACGTGTGGCCCTTT--GGTGTGCGCTGTGGGAGGTGCTGATGCTGT 760
Qy 2858 TAGGCGCGAGCCCTTTGGGCGAGCTCACCGAGCAGCAGTGCATGAGAACGCGGGGAGTT 2917
Db 761 GTAGGCCAGCC--TTTGGCAGCTACCCAGCAGGCGS----TCCCGAGACCCGGGGGAGT 814
Qy 2918 CTTCCGGGACCGAGGCGCG 2936
Db 815 CTTCCGGGACCGGCGCAGG 833

RESULT 15
LOCUS BE737252 957 bp mRNA EST 15-SEP-2000
DEFINITION 601305802F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3640248 5',
mRNA sequence.
ACCESSION BE737252
VERSION BE737252.1 GI:10151244
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 957)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM344 row: c column: 01
High quality sequence start: 22

```

High quality sequence stop: 699.
Location/Qualifiers
1. .957

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3640248"
/clone_lib="NIH_MCC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pORF7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."

BASE COUNT 137 a 301 c 318 g 201 t
ORIGIN

Query Match 16.5%; Score 652; DB 10; Length 957;

Best Local Similarity 93.5%; Pred. No. 6.5e-111;

Matches 757; Conservative 0; Mismatches 45; Indels 8; Gaps 7;

```
QY 1023 GTGGTGGGCTGGATGACTTTAGGAAGAGTCAGGAGCTGCGGCTCGCCAGGCTATGAC 1082
DB 8 GTGGGCTGCTGTTATGACTTTAGGAAGAGTCAGGAGCTGCGGCTCGCCAGGCTATGAC 67
QY 1083 TATGTGGGATGGAGCAACCAAGCTTCTCCAGTGCTATGAGATGGAGTTTGAGTTT 1142
DB 68 TATGTGGGATGGAGCAACCAAGCTTCTCCAGTGCTATGAGATGGAGTTTGAGTTT 127
QY 1143 GACCGGCTGAGGGCTTCCAGGCTATGAGGCTCCACTGTACACATGCACACGCTGGGA 1202
DB 128 GACCGGCTGAGGGCTTCCAGGCTATGAGGCTCCACTGTACACATGCACACGCTGGGA 187
QY 1203 GCCCGTCTGCTGGCGGGTGGAAATGTCGCTTCGGGCTGGCCCTGCCATGGCTGGGAG 1262
DB 188 GCCCGTCTGCTGGCGGGTGGAAATGTCGCTTCGGGCTGGCCCTGCCATGGCTGGGAG 247
QY 1263 GGGAGGCCATGCGCCCAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTC 1322
DB 248 GGGAGGCCATGCGCCCAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTC 307
QY 1323 TCAGTGCCCTTGGCGGCGTGGCTGCGCTTTCAGTGCCGCTGCTCTTTGCGGGG 1382
DB 308 TCAGTGCCCTTGGCGGCGTGGCTGCGCTTTCAGTGCCGCTGCTCTTTGCGGGG 367
QY 1383 CCCTGGTTACTTTCAGGAAATCTCCTTCATCTGATGTGGTGAACAATTCCTCTCCG 1442
DB 368 CCCTGGTTACTTTCAGGAAATCTCCTTCATCTGATGTGGTGAACAATTCCTCTCCG 427
QY 1443 GCATGGGAGGACACTTCCCGCCAGCCCTTGTGGCGGCTGGCCACCTCCACCAAC 1502
DB 428 GCATGGGAGGACACTTCCCGCCAGCCCTTGTGGCGGCTGGCCACCTCCACCAAC 487
QY 1503 TTGAGCAGCTTGGAGCTGGAGCCAGAGGCCAGAGCCGCGTGGCCAGGCCGATGGGA 1560
DB 488 TTGAGCAGCTTGGAGCTGGAGCCAGAGGCCAGAGGCCGCGTGGCCAGGCCGATGGGA 547
QY 1561 GCCCGACCGCCATCCTCATCGGCTGCGTG -GTGGCCATCATCTGCTGCTGCTCATC 1619
DB 548 GCCCGACCGCCATCCTCATCGGCTGCGTG -GTGGCCATCATCTGCTGCTGCTCATC 607
QY 1620 ATTGCCCTCATCTGCGGGCTGCACTGGCGAGGCTCCTCAGCAAGGCTGACGGAGG 1679
DB 608 ATTGCCCTCATCTGCGGG -TGGACTGCCGAGGCTCCTCAGCAGGCTGAACGGAGGG 666
QY 1680 GTGTTGGAAGAGAGCTGACGGTTACCTCTCTG -TCCCTGGGGACACTATCCTCATCAA 1738
DB 667 TTGTTGGAAGAGAGCTGACGGTTACCTCTCTGTTCCCTTGGGGACACTATTCTCATCA 726
```

```
QY 1739 CAACCGCCAGGTCTTAGAGAGCCACCGCGTACAGAGCCCGCGCTCGTGGAAATCC 1798
DB 727 AGACCGCCAGGTCTTAAGAGGAGCCACCGCTAC--AGGGCGCCCGCTCGTGGATACC 784
QY 1799 GCGCCACTCGGCTCCCTGTGTGTCCTCCCAATGG 1828
DB 785 GGCCCA-TCCGGTCCCTGTGTCGGCATTTG 813
```

Search completed: December 8, 2001, 20:10:20
Job time: 4419 sec

